

17. OERSTED'S DISCOVERY

Turning from these interesting but sketchy suggestions, I shall now present a fuller account of one of the greatest discoveries in the whole of the history of science – Oersted's discovery of electromagnetism¹⁷⁸. Although no discovery has gained so much unexplained praise, this most problematic discovery has never, to my knowledge, been satisfactorily explained. A striking example is Duhem's highly mathematical and very dry treatise on electricity in which he suddenly bursts into a eulogy of the importance of Oersted's discovery, which he describes as "the point of departure of all researches which constitute electrodynamics and electromagnetics"¹⁷⁹. This evaluation is correct; and yet it is puzzling that Duhem should say this⁵⁰, especially as he deals with the "point of departure" only at the end of his third and last volume.

The phenomenon which Oersted discovered, Oersted's effect, is easy to repeat. All that is needed are a compass, an electric battery, and a wire. When a direct current passes through a wire which is placed parallel to the needle of the compass, above or below it, the needle ceases to point north. If the wire is not exactly parallel to the needle, the effect will be smaller, but it will vanish only when the wire is very precisely perpendicular to the needle – in the east-west direction. The wire can be replaced by an electric torch, though the resistance of the glowing wire in the bulb will diminish the current and thus the effect.

Great experimenters, from Franklin to Davy, experimented with electricity and magnetism trying to find some connection between them; prizes were offered for the discoverer of this connection; Oersted himself thought and wrote about it for years before his own discovery. The editor who published the German translation of Oersted's first report of 1820, a Professor Gilbert, said it was clear that the discovery was purely accidental: what Oersted had failed for years to find while searching for it, he stumbled on during a public lecture! A popular myth developed according to which Oer-

sted made the discovery while trying to demonstrate the well-known fact that wires conducting electricity can glow.

Phillip Lenard has claimed that the fact that the discovery was made in a lecture can be explained by the fact that Oersted loved to lecture and did so frequently¹⁸⁰. It seems to me that the explanation provided by his pupil Hansteen is a better one. Hansteen reports¹⁸¹ that Oersted could not manipulate instruments – we know indeed from other sources that he was short-sighted and clumsy – and he used to ask members of his audience to experiment for him. In a recent publication R. C. Stauffer takes this as a personal insult to Oersted;¹⁷⁸ I take it to be a possible explanation of why he loved to lecture and prepare new experiments for his lectures. For my own part, I have a different explanation of the event. It is well known that soon after Oersted's discovery Ampère announced in his first lecture to the French Academy his discovery of the electro-magnetic solenoid, prior to observing it, and that Davy had some years before intended to decompose nitrogen and had planned to make the discovery in a public lecture. These can hardly be coincidences, and I conjecture that experimenters preferred to make discoveries before witnesses after explaining the ideas behind them, so as to forestall any claim that these discoveries were accidental.

The fact that Oersted's discovery occurred in the lecture room may or may not be significant; we cannot know this without knowing how much Oersted knew before entering the lecture room. The problem, then, is this: if he predicted his phenomenon, why did he not verify his prediction immediately? What difficulty did he encounter?

A Baconian might say that Oersted obviously made an essentially correct prediction, but was misled by a prejudice. "Obviously," because a Baconian cannot but identify impediments with prejudices. "We heard it stated by Faraday in one of his lectures at the Royal Institute," writes Bakewell¹⁸², "that when he was first connected with that institute, Davy and Young were frequently making experiments with the view of establishing the identity of electricity and magnetism, but that being misled by preconceived theories of the action of the force, they adopted nearly every conceivable mode but the simple one" adopted by Oersted. With this the efforts of such great men as Davy and Young are dismissed, and, allegedly, on Faraday's authority! But I want to find out what the actual prejudice was, and how Oersted overcame it, if he did indeed share it. All too often it is alleged that a person failed to discover what he was looking for until he laid aside his preconceived ideas, whereupon, lo and behold, he became successful. This is too schematic and preconceived; we need more details when discussing the history of a specific discovery. This theory, therefore, can simply be dismissed.

Another theory, also mentioned in the last section, is that the prediction

was incomplete; it was applied to our case by Hansteen, who has suggested¹⁸¹ that although Oersted, like everyone else, predicted the discovery in general, he did not predict it in detail; he got the detail by accident rather than by prediction. The important detail is that the electromagnetic force seems to be circular or rotational or transversal (see section 5, above); from this it follows that the strongest deflection of the needle would occur when the current is placed parallel to it, in the south-north direction. This, Hansteen states, Oersted did not know, as is shown by the fact that he first placed the wire in the east-west direction.

Hansteen's story is quite satisfying to a historian. That Oersted did not know the fact that the force is rotational, Hansteen shows by using historical evidence – the fact that Oersted used to place the wire in the east-west direction. Hansteen describes the accident by which this fact was discovered, and his description is so vivid, detailed, and natural, that we must accept it, if not as evidence by an eye-witness then at least as a serious hypothesis, or as a story which he had received from the horse's mouth and which was probably common knowledge in Copenhagen's scientific circles. "Once, after the end of his lecture, as he had used a strong galvanic battery for other experiments, he said [to his assistant]: 'Let us now once, as the battery is in activity, try to place the wire parallel with the needle'; as this was done, he was quite struck with perplexity by seeing the needle making a great oscillation . . ." All one needs to remember is that in Oersted's time an electric battery was operative for a few minutes only, and one will vividly see before one's eyes the eager experimenter snatching the opportunity after a lecture and just trying whatever occurred to him first, and luckily hitting upon the discovery.

R. C. Stauffer has argued that Hansteen was not an eye-witness by reproducing a personal letter from Oersted to Hansteen¹⁸³ (1820) in which the discovery is discussed. He concludes that Hansteen's evidence is "worthless" as a historical document though his story is "very plausible." Later he claims on the strength of a few reports of the discovery written by Oersted himself that the discovery was not accidental after all.

For my own part, I do not accept Hansteen's story, yet I do not think that the factual information it contains can be discarded without sufficient reason, especially since he was a close friend of Oersted, and since Oersted's own reports are sketchy and extremely difficult to understand. My disagreements with Hansteen are three. First, it is untrue that, like many people, Oersted had expected some relation between electricity and magnetism; significantly, his theory was very different from other theories, and it led him to introduce the electric current into the investigation. Secondly, the current itself was not there accidentally nor was its role predicted by anyone but Oersted. Thirdly, as to the direction of the current, it is unlikely that Oer-

sted would have persistently placed the current in the east-west direction, and then, by accident, in the south-north direction. Hitting on any of these two out of infinitely many directions by sheer accident is most unlikely.

Our starting point is the theory which I have already mentioned (section 5): only likes interact; and electricity and magnetism are not likes; hence they cannot interact. The idea that only likes interact was universally accepted until much later. This explains why those who thought that electricity and magnetism *do* interact spoke of the *identity* of the two. Those like Ampère who rejected this *identity* rejected the possibility of interaction as absurd.

This is all highly metaphysical, yet there is nothing else to begin with. Whittaker, ignoring this metaphysical discussion (he was an inductivist) and trying to find some continuity here (he accepted the continuity theory) refers¹⁸⁴ to some reports about occurrences of interaction between electricity and magnetism which were known before 1820. But if these reports were acceptable, Ampère (and many other empiricists) would have been anti-empiricists, and the academies which offered prizes for those who could find the interaction, associations of fools. Indeed, Whittaker himself admits in a way that the reports were unacceptable. All the known cases of interaction between electricity and magnetism – the most important of which was the effect of lightning on compasses – were unrepeatable, and unrepeatable experiments do not count. (We know of many experiments which are entirely ignored, although they would count as great discoveries if they had been repeatable.) Even Oersted himself, who drew much encouragement from the unrepeatable experiments of his friend Ritter, was somewhat ashamed of the fact that so great an experimenter as Ritter should have published unrepeatable experiments. We shall return to this later; here we must conclude that there was no proper experimental background to Oersted's discovery.

The doctrine that only likes interact was finally shaken by Nicholson and Carlisle, who discovered in 1800 the phenomena of electrochemistry; namely, the fact that electric forces can oppose and overcome the forces of chemical affinity. Electrochemistry indicates that electric forces and chemical forces are identical, and this indeed was Davy's and Berzelius' guiding idea. But this idea does not suffice; electric forces belong to electric matter and chemical forces to gross matter, so that they must be different and therefore cannot interact! Davy and Berzelius did try to cope with this problem, but, I think, rather unsuccessfully. Oersted, on the other hand, tackled the problem-situation from a new and wider angle. Chemical affinity and electricity, he thought, are neither identical nor essentially different; they are different manifestations of essentially one and the same force; all allegedly different forces are different manifestations of one and the same primordial force.

The primordial force can change its manifestations, and the task of the experimenter is to find the conditions under which such changes take place. This is Oersted's own original idea¹⁸⁵, and one which guided the whole of his scientific and philosophical career from the time he started thinking about the problems posed by electrochemistry to the end. The greatest objection to his doctrine, he thought, was the seeming impossibility of converting electricity into magnetism. Accordingly, he looked for a way around this difficulty. In the book about chemistry he published in 1813, he wrote that perhaps galvanic electricity interacts with magnets more easily than frictional electricity. In the subsidiary reports¹⁸⁶ about his discovery he quoted himself on this point as evidence that his discovery had been predicted.

I find this argument amazingly unconvincing. Other philosophers had made this prediction, and they had even tried, although unsuccessfully, to confirm it; this has been very well known. The intellectual background to the discovery makes it easier to understand why Oersted persisted in his attempts after other philosophers had abandoned their own. But we have already come a long way from the casual remark of 1813 – which, by the way, is entirely false – to the discovery in 1820.

The next step concerns the introduction of the (galvanic) current. Again Oersted's own report is cryptic. My attempt to reconstruct his thought makes use of his theory of the current. He had noticed long beforehand that since, according to his own view, there is only a variety of manifestations of force and not a variety of matters, he could not accept the orthodox theory of electric currents or discharges as (literally) currents or flows of electric matter (in the fluid state). He wanted to explain this phenomenon. Now in electric currents or discharges, as he knew, electric forces transform into heat and light. So this is what currents are – transformations of forces; their mechanism, he assumed, was a kind of state of disequilibrium of conflicting electric forces along the wire which unites the electric poles.

It is now clear why Oersted's speculations were not taken seriously; why indeed, they and their originator were viewed with hostility almost to the point of public scandal. Not only was the prevailing (Baconian) scientific climate opposed to speculation; Oersted's speculations were opposed to orthodox Newtonian speculations and akin to the German Romantic speculations accepted in circles which were inimical to experimental philosophy¹⁸⁷.

Returning to Oersted's train of thought, we have seen that he had thought that in an electric discharge the electric force is transformed into other kinds of force; namely, heat and light. And now (in 1813 or perhaps somewhat later) he came to think that if the current is sufficiently strong it might also turn electricity into magnetism. In one subsequent report he writes: "somewhat inconsistently, I expected the predicted effect particularly from the discharge of a large electric battery and moreover only hoped

for a weak magnetic effect.” I do not understand what “somewhat inconsistently” means here unless it means “erroneously”, and I think one can explain his mistaken expectations concerning the magnitude of the effect as well as his mistaken expectation that the effect would take place only while the current generates both heat and light. He might have viewed the current like this: when a large concentration of electric forces in disequilibrium (i.e., a current) occurs, they are transformed into heat; when a larger concentration occurs (the wire is thinner or the current stronger) a small residuum is transformed into light as well. So, by a kind of extrapolation – perhaps when a still larger concentration occurs – the transformation of a smaller residuum into magnetism may take place. And the fact that the residuum was so small would explain, incidentally, why people have not noticed it before.

Moreover, Oersted thought that the glowing wire was a magnet, but he did not know any more about it, and in particular he did not know where its poles lay. We know it has no poles but has circular rather than central – attractive or repulsive – forces. Oersted not only did not believe that the force is circular; he definitely thought that it was central. We have ample evidence of this from his own writings, and we might have expected it in any case since he was at once a kind of Newtonian and a kind of Kantian. Now, if one has a long weak magnet, if one does not know where its poles lie or which is north and which south, and if one wishes it to interact with a compass, some knowledge of Newton’s theory of force will tell one to place the magnet in the east-west direction. One does so and sees no result. Hence one appears to have made a mistake. One concludes that either (a) the long weak magnet is weaker than thought, or (b) that it is not a magnet after all, or else (c) that the Newtonian hypothesis concerning forces is false.

Remembering that Oersted was a kind of Newtonian and apriorist, one sees at once that he would have to reject corrections (b) and (c) without much thought and accept correction (a), which entails that he would have to repeat the experiment with a still stronger battery in order to get a result.

At least one interesting intermediate step occurred while he was trying to obtain stronger and stronger batteries. The intermediate step was to place the wire not at all above the compass but at its side, perpendicular to the plane in which it lay: Oersted took a hint from the dubious reports that lightning – which Franklin had identified as currents – affected compasses. It was this idea which occurred to him just before that lecture in March or April 1820, and which he tried out during the lecture. He did get some results, but they were still irregular; his audience remained totally unimpressed. (This intermediate step shows how reluctant he was to place

wire in the south-north direction, a fact which I have explained by reference to Newton's theory of force). Remembering Ritter's irregular results, he accepted his audience's judgment and refrained from publishing; indeed, he left the whole business unattended for three months. Returning then to his experiments, he constantly increased the size of his battery. Thus, he was still trying to stick to his Newtonianism, or else perhaps he still had not thought of criticizing it. His line of attack was still on traditional lines – or at least it can be so explained, as I shall now argue.

Having had an irregular result from using the wire in the place of a lightning flash, he tried to find out where the pole lies in the wire. He tried to find the pole in the wire, like a needle in a haystack, by bending it in various places into a V-shape. The results of this experiment are indeed regular. So, having obtained a regular though intermediate result, he decided, I conjecture, to go on trying to locate the poles in public before an audience of highly distinguished gentlemen, while demonstrating to them that he had got some regular result. Although he did make his point, he could not explain it: he found no poles. The poor lecturer had no option but to declare the lecture finished. With some sense of excitement and with more than some sense of disappointment and embarrassment his distinguished guests rose and made for the door; but they were stopped before they got out. For in the mean time the discovery had been made. From a deep sense of despair Oersted quickly, but very cruelly, took stock of his activities, and tried to explain his strange failure. No doubt it was useless to spend one's life (in a Kafka-like fashion) building ever bigger batteries. The error lay somewhere else. Was he mistaken in believing that the conducting wire was a magnet? Impossible. What other systematic error did he commit? I doubt that he could there and then have supposed that Newton's theory of force was possibly false. What is clear is that he quickly suggested that his systematic error from the beginning of his series of experiments to the end lay in placing the wire in the wrong place. He may have thought so either without asking why, or while (erroneously) thinking that this early deduction of the mistaken east-west direction of the wire from the correct Newtonian theory of force was invalid, or else while suggesting that the Newtonian theory of force was incorrect. Anyhow, he was in a frantic hurry; it was now or never – before the distinguished guests had reached the door. Quickly he placed the wire in the north-south directions, possibly just to get as far away from the early systematic error as possible. And then he gasped; he saw at once how much more important his discovery was than he had ever hoped.

This concludes my reconstruction except for a brief discussion of a few remaining difficulties.

First, the problem of the magnitude of the effect, and hence also the

problem of the erroneous use of a glowing wire, is practically irrelevant. Why, then, does Oersted so stress these red herrings? I do not know, but feel that he had to explain the delay from 1813 to 1820 and could not very well blame the real obstacle, the Newtonian hypothesis that all forces are central. So perhaps he blamed the small and irregular character of the effect. But there is more to it; admittedly, the need for a big battery, which he retained to the last, but which entirely vanished at the moment of the discovery of the circularity of the force, is quite eloquent evidence that Newtonianism was an obstacle; yet had he started with a strong battery and a thick wire, and placed the wire (erroneously) in the east-west direction, he would have had a nearly fifty percent chance to observe the needle making a full about-turn! Thus, we see that there were several errors interplaying, and that Oersted preferred to name his own – the use of thin wires – as the chief culprit.

The next difficulty is this. Oersted claimed that the discovery was made early in 1820, while I claim that it was made in July 1820, three months later. Here, however, there is no difference about details, only about their evaluation¹⁸⁸; and we should remember that Oersted was anxious to refute claims that the discovery was accidental as well as claims to priority on the part of some other person.

There remains one final objection: Hansteen says that the discovery of the law of electromagnetism was found (immediately?) *after a lecture*¹⁸⁹, while I claim that *during* the wintertime lecture Oersted found only irregular results. But I have the whole of Oersted's story to support me. Possibly Hansteen confused the lecture in winter with the discovery in July. I prefer to think, however, that he was referring not to a formal University course but to a demonstration before the very distinguished audience. That Oersted would have made a major step in public twice seems rather unlikely, of course. But Hansteen explains it entirely satisfactorily, I think, when he says that Oersted usually experimented during his lectures, and asked members of the audience to act as his assistants.